Attorney Docket No. 50269-0516 *Patent*

UNITED STATES PATENT APPLICATION

FOR

CONTENT-BASED BILLING

INVENTORS:

PEI-YUAN ZHOU
IVRY SEMEL
SAILENDRA K. PADALA
PHILIPPE LE ROHELEC

PREPARED BY:

HICKMAN PALERMO TRUONG & BECKER, LLP
1600 WILLOW STREET
SAN JOSE, CALIFORNIA 95125
(408) 414-1080

CONTENT BASED BILLING

PRIORITY CLAIM AND CROSS-REFERENCE TO RELATED APPLICATIONS

This application is related to and claims domestic priority from prior U.S. Provisional Application Serial Number 60/269,699 (attorney docket number 50269-0514), filed on February 17, 2001 entitled "Content-Based Billing and Header Based Network API", by Michael M. Tso, Pei-Yuan Zhou, Ivry Semel, Sailendrak Padala, and Philippe Le Rohelec, the entire disclosure of which is hereby incorporated by reference as if fully set forth herein.

FIELD OF THE INVENTION

The present invention relates to applications on the World Wide Web and more specifically to content-based billing on the World Wide Web.

BACKGROUND OF THE INVENTION

Most people that access electronic content do so through an access provider. In the most common scenario, the device used to access electronic content is a desktop computer, the source of the electronic content is a server on the Internet, and the access provider is an Internet Service Provider (ISP). In most cases, an ISP provides one or more devices, referred to herein as access servers, through which users can connect to the Internet. The access servers simply act as pipes through which all messages to and from the user pass.

Typically, the ISP knows little about a user's use of the Internet beyond the times at which the user logs on and logs off, and the number of bytes that pass through the access server in the user's session. Consequently, the ISP has little flexibility with respect to how it will bill for its service. Generally, most ISPs that do not provide wireless connections charge

a flat-rate subscription fee, or a time-based usage fee. Access providers that provide access to wireless devices often bill based on "airtime", or the number of bytes retrieved by the user.

For a variety of reasons, it may be desirable for access providers to bill based on other factors. For example, the Internet session of a user of a mobile device, such as an Internet-enabled cell phone or PDA, may include a high percentage of idle time. If the user was charged strictly based on time, the cost of the session may be out of proportion to the service received by the user. As another example, the access provider may want to provide some services, such as advertising, that the user would be able to access without being billed. Conversely, the access provider may want to charge premium rates for the use of other services.

Based on the foregoing, it is clearly desirable to provide systems and techniques that allow access providers to have billing policies that are more flexible than are currently available.

SUMMARY OF THE INVENTION

Techniques are provided for intercepting, modifying, and re-transmitting a request between various entities engaged in content-based billing. Various techniques, involving devices such as HTTP proxy servers, wireless protocol gateways, routers, and level 4 switches, may be used for intercepting a request for content. The request maybe modified based on numerous factors such as profile information, levels of service, and requested content. The entities that engage in content-based billing include but are not limited to users (including mobile users), content providers, and access providers (including HTTP proxy servers). The ability to bill users and/or content providers based on the content that is exchanged between the parties enables the use of billing policies that are much more flexible than those currently in use.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

Figure 1 is a block diagram of a system that allows flexible billing schemes according to an embodiment of the invention; and

Figure 2 is a block diagram of a computer system on which embodiments of the invention may be implemented.

DETAILED DESCRIPTION OF THE INVENTION

A method and apparatus for content-based billing is described. In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to avoid unnecessarily obscuring the present invention.

FUNCTIONAL OVERVIEW

FIG. 1 is a block diagram of a system that allows flexible billing schemes according to an embodiment of the invention. Referring to FIG. 1, users 170 access services provided by content servers 160 through an access server 100. However, rather than simply act as a pipe for messages, the access server 100 interacts with an HTTP proxy server 120. The HTTP proxy server 120 inspects messages that pass between each user 170 and the content servers 160. The proxy server 120 generates a log 130 to indicate which users (170a, 170b, or 170c) are accessing which services. The access logs 130 are provided to an aggregation engine 140 that places the access information into the format required by the billing system 150. The access information is then fed to the billing system 150, and the users 170 are billed, at least in part, based on the content that they access through the access server 100.

Under certain circumstances, it may be desirable to pre-bill the user 170a prior to providing the user 170a with requested content. Under these circumstances, prior to providing the content to the user 170a, the HTTP proxy 120 may access a source to see if the user 170a has sufficient funds. That source may be, for example, a user profile 110 associated with the user 170a, or the billing system 150 itself. If the user 170a does not have sufficient funds, then the HTTP proxy 120 does not deliver the requested content to the user 170a. If the user 170a does have sufficient funds, then the HTTP proxy 120 causes the funds

to be decremented based on the fee associated with the content, and then forwards the content to the user 170a.

SERVICES

An access provider (as represented among other things by access server 100 and HTTP proxy 120) may provide services beyond simply connecting users to a network, such as the Internet. For example, access providers for wireless devices often provide a bundle of services from which their customers may choose. The mobile device displays a menu, where each menu item corresponds to a service. The service may be for (1) content provided by the access provider itself, (2) content provided by third parties (e.g., represented by one or more content servers 160), or (3) some combination of the two.

For the purpose of illustration, it shall be assumed that the content associated with services comes from content servers 160 that are separate from the access server 100. However, the techniques described herein are equally applicable when the access provider is also the provider of the content.

USER IDENTIFICATION

Once a user 170a has successfully logged in to the access server 100, the access server 100 typically assigns the user 170a a unique address, and all messages sent by the user 170a during that session include the unique address. For the purpose of explanation, it shall be assumed that the unique address assigned to devices is an IP address, though the actual type of address may vary from implementation to implementation.

The HTTP server 120 sees the IP address in each message that it intercepts, and must be able to determine the user 170a that has been assigned the IP address. According to one embodiment, the user ID-to-IP address mapping for user 170a is communicated from the access server 100 to the HTTP proxy 120 at the time each user (e.g., 170a, 170b, and 170c)

initiates a session. Various techniques may be used to communicate this information, including:

- the access server 100 sending to the HTTP proxy 120 a packet with the information using a particular protocol supported by the HTTP proxy 120, such as the RADIUS protocol;

- the access server 120 sending to a "protocol proxy" a packet with the information using a particular protocol, such as the RADIUS protocol, and then having the protocol proxy send the information from the packet to the HTTP proxy 120 using a different protocol that is supported by the HTTP proxy 120; and

- the HTTP proxy 120 publishes an API to the access server 100, and the access server 100 makes a call through the API to communicate the information to the HTTP proxy 120.

Preferably, the mechanism employed allows the access provider to use a generic HTTP proxy, thereby avoiding the programming burden that a custom HTTP proxy server would impose.

MESSAGE INTERCEPTION

The HTTP proxy 120 shown in FIG. 1 intercepts messages between user devices such as the devices depicted for users 170 and content servers 160. After intercepting the messages, the HTTP server 120 inspects the messages to see what content the user 170a is requesting/receiving. If the request/receipt of the content has billing implications under the access provider's billing scheme, then the HTTP proxy 120 generates a log 130 to indicate that the user 170a requested/received the content. In addition to generating a log 130 to indicate which services/content the user 170a requested/received, the HTTP server 120 may handle the message in a different manner than simply forwarding the message. Various

examples of how the HTTP proxy server 120 may handle messages in particular situations shall be described hereafter.

The illustrated embodiment uses an HTTP proxy 120 to perform the message interception. However, using HTTP proxy 120 is only one of many techniques for performing the message interception. Other known techniques for intercepting messages involve the use of wireless protocol gateways, routers, and level 4 switches. The present invention is not limited to any particular technique or mechanisms for intercepting messages. Preferably, the mechanism used to intercept messages is capable of both intercepting messages that are in transit, and of reinserting messages (often after some modification) back into transit. It is also preferable that the mechanism be capable of intercepting both messages in both directions: from users 170 to servers 160, and from servers 160 to users 170.

USER PROFILES

Different users may register for different services, or different classes of the same service. According to one embodiment, the access server 100 maintains user profiles (of which user profile 110 is an example) that indicate, among other things, the services for which each user (e.g., any one of the users 170) has registered.

Using the techniques described herein, the access provider is able to establish flexible and sophisticated billing schemes. Part of that flexibility results from the ability to store user-specific billing data in the user profiles 110. For example, if a given service supports different classes of service, then the user profile 110 of a user 170a will also indicate the class of the service for which the user 170a is registered. The class of service reflected in the profile 110 may be used to determine how much the user 170a is to be billed when the user 170a accesses the service.

The access providers may also have a classification system for users 170, where certain "gold" members are charged certain rates for services, while "silver" members are charged different rates. Rate plans (described in more detail hereinafter) are one way of providing these different rates. The user's profile 110 may indicate the user class to which the user 170a belongs.

REDIRECTING REQUESTS

According to one aspect of the invention, the HTTP proxy 120 may redirect an intercepted request based on the profile of the user that sent the request. For example, a user 170a may send a message requesting a stock quote from a stock quote service. If the user 170a belongs to the "gold" class for the service, then user 170a is allowed to access a real-time stock quote. If the user 170a belongs to the "silver" class for the service, then the user 170a is allowed to access a quote that is 10 minutes old. In response to determining that the request is for a stock quote, the HTTP proxy 120 reads the profile 110 of the user 170a to determine the class of service for which the user 170a is subscribed, and redirects the request to the appropriate source. For example, the HTTP proxy 120 may direct the request to one URL if the user 170a belongs to the gold class, and to a different URL if the user 170a belongs to the silver class.

By putting the intelligence to redirect messages to appropriate destinations on the server side, greater flexibility can be achieved without adding complexity to the user devices. For example, the fact that a user 170a belongs to particular classes for particular services is completely transparent to the user's device. This is particularly important when the user devices are relatively small and unsophisticated mobile devices.

Redirecting a request is just one example of how the HTTP proxy 120 may alter and/or transform a request based on information in the user's profile 110, and the identity of the content that is being requested and/or received. As another example, the HTTP proxy 120 may determine that the user 170c is requesting a service that requires certain information about the user 170c, such as the current location of the mobile user 170c. The HTTP proxy 120 may, transparent to the user 170c, insert into the request the additional content required by the service prior to forwarding the request to the content server 160a for that service.

The general steps are:

- intercepting the request
- identifying the user 170c that submitted the request
- identifying the profile 110 that corresponds to the user 170c
- modifying the request based on the user's profile 110
- retransmitting the modified request

UPDATING USER PROFILES

According to one aspect of the invention, the profile 110 of a user 170a may be updated based on the content requested/received by the user 170a. For example, the user profile 110 may indicate that the user 170a is allowed to access a particular service three times a month for free, and then is charged for any additional accesses. Under these circumstances, the user's profile 110 may include a number that indicates how many times the user 170a has accessed the service during the current month. Every time the user 170a accesses the service, the number is incremented until it reaches three. Once it has reached three, then rather than increment the number in response to the user 170a requesting the service, the HTTP proxy 120 generates a log 130 that, when supplied to the billing system

150, will cause the user 170a to be billed. The number would then be reset at the beginning of each month.

The above example is simply one of many situations in which it is desirable to modify the profile 110 of a user 170a in response to the content that is being requested/received by the user 170a. Other examples include decrementing an account balance stored in the profile 110 when the user 170a accesses services that have fees, marking a user 170a for participation in a prize drawing in response to reward the user 170a for looking at advertisements, etc.

AUTHORIZATION

Most access providers have a mechanism for determining that users are allowed to access their services. For example, when a user 170a attempts to establish a connection to the Internet through an ISP, the ISP typically requests the user ID and password of the user 170a. If the user ID/password combination matches that of an authorized user 170a, then the requested connection is established. Otherwise, the user 170a is denied the connection.

According to one aspect of the invention, the access provider's ability to authorize user action is extended beyond the mere ability to grant or deny a connection. Rather, since the HTTP proxy 120 is monitoring the content that is requested/received by a user 170a, the access provider is able to grant or deny access to specific content. For example, the user 170a may have subscribed to a certain class of a particular service. If the user 170a requests content that is not available to that certain class, then the HTTP proxy 120 can send the user 170a an appropriate message, and prevent the user's request from being forwarded to the content server 160a.

CONTENT PROVIDER PROFILES

According to one aspect of the invention, the techniques described above for handling users 170 may be applied in a similar manner to content providers (as represented by one or more content servers such as 160a, 160b, 160c, and 160d). For example, the HTTP proxy 120 may store content provider profiles, and examine the profiles to determine how to process content received from the content providers. For example, just as the users 170 are charged for accessing certain content, content providers may be charged for providing certain content.

For example, the access provider may provide a service that allows users 170 to make purchases from a particular content provider. The access provider may enter an agreement with the content provider that give the access provider a commission on purchases made by its users 170. Under these circumstances, the HTTP proxy 120 may be configured to generate a log 130 when it detects that the content provided by the content provider indicates a purchase by a user 170a. This log information 130 can be fed to the billing system 150 by the aggregation engine 140 to bill the content provider for the commission.

Thus, during any transaction conducted through messages that are intercepted by the HTTP proxy 120, the HTTP proxy 120 may inspect the profiles for any of the participants (such as 170a, 170b, 170c, 160a, 160b, 160c, 160d) in the transaction, modify the profiles for any of the participants in the transaction, and bill any of the participants in the transaction.

PROCESSING LOGS

In the embodiment shown in FIG. 1, an aggregation engine 140 processes the logs 130 generated by the HTTP proxy 120 and feeds the data into the billing system 150. In one embodiment, a single aggregation engine 140 may be used to process the logs of many HTTP

proxies 120. For example, one access provider may have multiple access servers 100 and multiple HTTP proxies 120, but a single billing system 150. The logs 130 from all HTTP proxies120 may be processed by a single aggregation engine 140, which sends the results to the billing system 150.

COMPONENT COMMUNICATION

In the techniques described above, communication occurs between various distinct components. For example, the access server 120 communicates the user ID-to-IP address mapping information to the HTTP proxy 120 when a user 170a starts a session, the aggregate engine 140 communicates with the billing system 150, etc. Preferably, these communications take place using a standardized protocol established for these communications.

SAMPLE SCENARIOS

Scenario 1:

A user 170a requests a service.

The HTTP proxy server 120 intercepts the request.

The HTTP proxy server 120 determines the user ID (for user 170a) based on the IP address.

The HTTP proxy server 120 determines, from the user's profile 110, the service class of that service for that user 170a.

The HTTP proxy server 120 modifies the request based on the service class and transmits the modified request.

The HTTP proxy server 120 intercepts the response back from the content server 160a.

The HTTP proxy server 120 determines the content provider ID based on the URL associated with the content.

The HTTP proxy server 120 determines, from the user's profile 110 and/or the content provider's profile, that the user 170a is to pre-pay fifty cents to receive the content.

The HTTP proxy server 120 sends a message to the user 170a asking the user 170a if the user 170a authorizes payment.

In response to receiving authorization, the HTTP proxy server 120 does the following:

forwards the content to the user 170a;

obtains the price of the service that the user 170a requested from the billing system 150; and

modifies the user's profile 110 to deduct fifty cents from an account balance indicated in the user's profile 110.

Scenario 2:

A user 170a requests a service.

The HTTP proxy server 120 intercepts the request.

The HTTP proxy server 120 determines the user ID based on the IP address.

The HTTP proxy server 120 obtains the price of the service that the user 170a requested from the billing system 150.

The HTTP proxy server 120 sends a message to the user 170a authorizing the payment of the fee;

The HTTP proxy server 120 receives a message from the user 170a indicating that the payment is not authorized.

The HTTP proxy server 120 does not forward the request to the content server 160a.

Scenario 3:

A user 170a requests a service.

The HTTP proxy server 120 intercepts the request.

The HTTP proxy server 120 determines the user ID based on the IP address.

The HTTP proxy server 120 obtains the price of the service that the user 170a requested from the billing system 150.

The HTTP proxy server 120 forwards the request to the content server 160a.

The HTTP proxy server 120 intercepts the content from the content server 160a.

The HTTP proxy server 120 inspects the profile of the content provider to determine that the content provider is to be billed for providing that content.

The HTTP proxy server 120 provides the content to the user 170a.

The HTTP proxy server 120 generates a log 130 indicating that the user 170a requested the service, and the amount the user 170a is to be billed;

The HTTP proxy server 120 generates a log 130 indicating that the content provider is to be billed.

The aggregate engine 140 sends the logs 130 to the billing system 150.

Bills are sent to the user 170a and the content provider

AN OPERATIONAL EXAMPLE

A user 170a subscribes to a service, such as ATT wireless for voice capabilities, and in so doing the access provider (e.g., ATT) creates an account for the user 170a. Then sometime later, the user 170a decides to add content billing capability for accessing data

services via URLs in which case the user 170a buys a deal. In response to the deal purchase, the access provider creates a user profile 110 that identifies the deals purchased by the user 170a. Each deal maps to a set of URLs. The URLs correspond to products. Thus, the deals purchased by a user 170a determine the set of URLs the user 170a may access, and consequently the products that are available to the user 170a.

When a user 170a requests a product by accessing a particular URL, the user profile 110 is inspected to determine whether the user 170a is authorized to access that URL. Once the product has been sent to the user 170a, the access provider sends to the billing system 150 a record that contains the details of the transaction. Within the billing system 150, the record is compared with one or more rate plans to determine a per-use charge for that transaction. The per-use charge thus determined may be in addition to a flat rate charged for the deals purchased by the user 170a.

The rate plans are price lists that may take into account a variety of factors beyond the identity of the product purchased, such as the time of day, the volume of data transferred, the time consumed in the transaction, etc. The rate plans that apply to a particular user may vary based on the deals purchased by the user 170a.

HARDWARE OVERVIEW

Figure 2 is a block diagram that illustrates a computer system 200 upon which an embodiment of the invention may be implemented. Computer system 200 includes a bus 202 or other communication mechanism for communicating information, and a processor 204 coupled with bus 202 for processing information. Computer system 200 also includes a main memory 206, such as a random access memory (RAM) or other dynamic storage device, coupled to bus 202 for storing information and instructions to be executed by processor 204. Main memory 206 also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor 204. Computer

system 200 further includes a read only memory (ROM) 208 or other static storage device coupled to bus 202 for storing static information and instructions for processor 204. A storage device 210, such as a magnetic disk or optical disk, is provided and coupled to bus 202 for storing information and instructions.

Computer system 200 may be coupled via bus 202 to a display 212, such as a cathode ray tube (CRT), for displaying information to a computer user. An input device 214, including alphanumeric and other keys, is coupled to bus 202 for communicating information and command selections to processor 204. Another type of user input device is cursor control 216, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to processor 204 and for controlling cursor movement on display 212. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), that allows the device to specify positions in a plane.

The invention is related to the use of computer system 200 for implementing the techniques described herein. According to one embodiment of the invention, those techniques are performed by computer system 200 in response to processor 204 executing one or more sequences of one or more instructions contained in main memory 206. Such instructions may be read into main memory 206 from another computer-readable medium, such as storage device 210. Execution of the sequences of instructions contained in main memory 206 causes processor 204 to perform the process steps described herein. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the invention. Thus, embodiments of the invention are not limited to any specific combination of hardware circuitry and software.

The term "computer-readable medium" as used herein refers to any medium that participates in providing instructions to processor 204 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example, optical or magnetic disks, such as storage

device 210. Volatile media includes dynamic memory, such as main memory 206. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise bus 202. Transmission media can also take the form of acoustic or light waves, such as those generated during radio-wave and infra-red data communications.

Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, or any other magnetic medium, a CD-ROM, any other optical medium, punchcards, papertape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described hereinafter, or any other medium from which a computer can read.

Various forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to processor 204 for execution. For example, the instructions may initially be carried on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to computer system 200 can receive the data on the telephone line and use an infra-red transmitter to convert the data to an infra-red signal. An infra-red detector can receive the data carried in the infra-red signal and appropriate circuitry can place the data on bus 202. Bus 202 carries the data to main memory 206, from which processor 204 retrieves and executes the instructions. The instructions received by main memory 206 may optionally be stored on storage device 210 either before or after execution by processor 204.

Computer system 200 also includes a communication interface 218 coupled to bus 202. Communication interface 218 provides a two-way data communication coupling to a network link 220 that is connected to a local network 222. For example, communication interface 218 may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, communication interface 218 may be a local area network (LAN) card to

provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation, communication interface 218 sends and receives electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

Network link 220 typically provides data communication through one or more networks to other data devices. For example, network link 220 may provide a connection through local network 222 to a host computer 224 or to data equipment operated by an Internet Service Provider (ISP) 226. ISP 226 in turn provides data communication services through the world wide packet data communication network now commonly referred to as the "Internet" 228. Local network 222 and Internet 228 both use electrical, electromagnetic or optical signals that carry digital data streams. The signals through the various networks and the signals on network link 220 and through communication interface 218, which carry the digital data to and from computer system 200, are exemplary forms of carrier waves transporting the information.

Computer system 200 can send messages and receive data, including program code, through the network(s), network link 220 and communication interface 218. In the Internet example, a server 230 might transmit a requested code for an application program through Internet 228, ISP 226, local network 222 and communication interface 218.

The received code may be executed by processor 204 as it is received, and/or stored in storage device 210, or other non-volatile storage for later execution. In this manner, computer system 200 may obtain application code in the form of a carrier wave.

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.